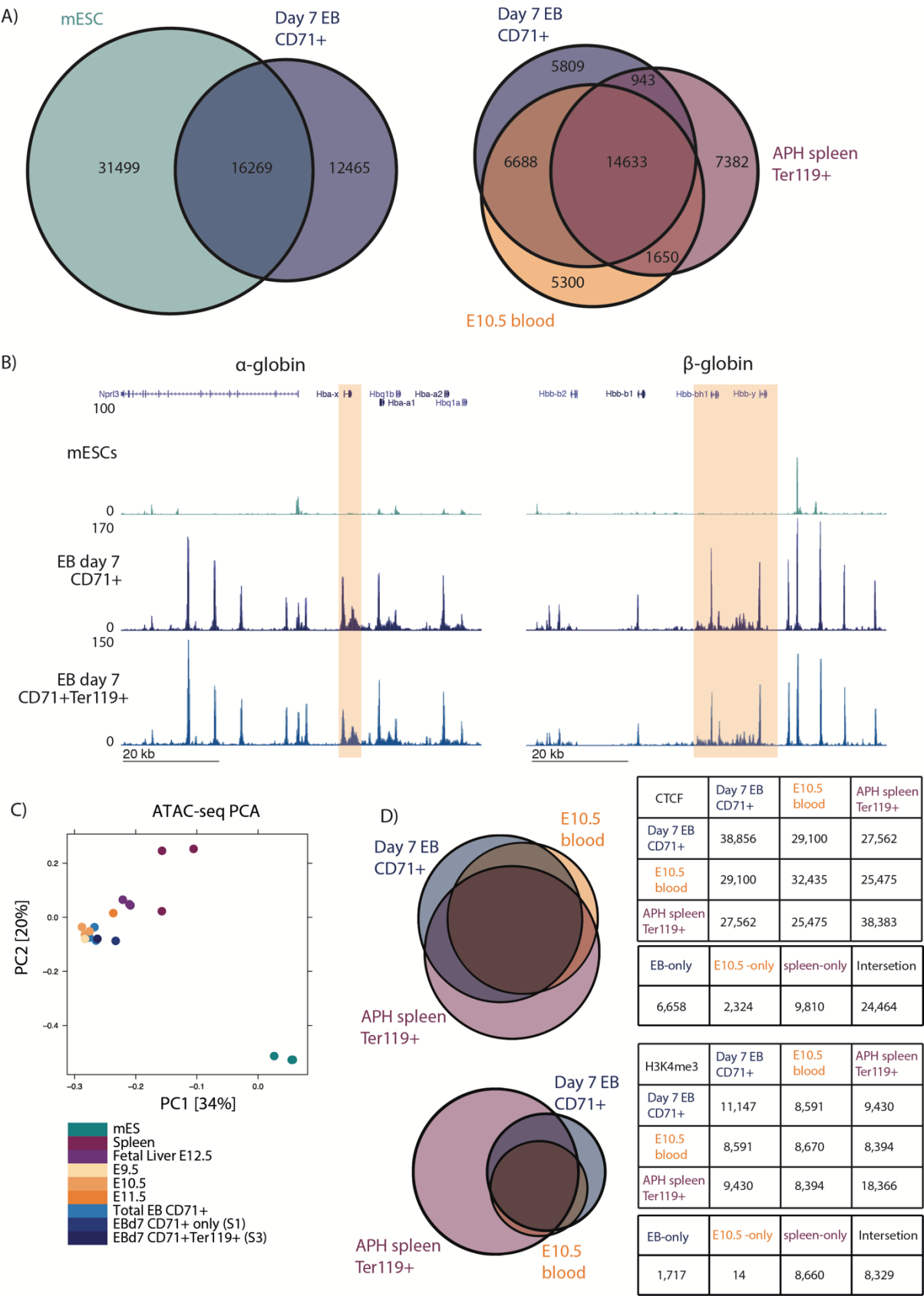


S1 Figure



**S1 Fig:** More detailed ATAC-Seq analyses of CD71 sub-populations (CD71+ S1, CD71+/Ter119+ S3) and CTCF and H3K4me3 ChIP-Seq data confirm the primitive nature of these populations.

**A)** Venn diagrams of overlapping accessible regions between tissue datasets. Peaks were included in the analysis (if present) in data from three biological replicate tracks. Left: mESCs vs EB-derived cells; Right: EB-derived, primitive and definitive erythroid cells.

**B)** RPKM-normalised ATAC-seq track averaged for three replicates of double positive CD71+Ter119+ cells compared to data from total CD71+ populations, both derived from day 7 EBs.

**C)** PCA plot of genome-wide ATAC-seq peaks, as in Figure 2B, with addition of primitive E9.5, E10.5 and E10.5 blood, the EB-derived CD71+ (S1) and CD71+/Ter119+ (S3) populations.

**D)** Venn diagrams of overlapping ChIP-seq peaks between tissue datasets. Peaks included in the analysis are from three biological replicate tracks for APH Ter119+ spleen and E10.5 blood and duplicates from EB day 7 CD71+ population. Upper diagram: CTCF ChIP-seq peaks with tables containing the detailed dataset from all three cell types. Lower diagram: H3K4me3 ChIP-seq peaks presented similarly as above.